INTERNATIONAL CENTRE FOR NUMERICAL METHODS IN ENGINEERING

Comparison of some finite element methods for solving the diffusion-convection-reaction equation

Ramon Codina

Publication CIMNE Nº 101, 1996

COMPARISON OF SOME FINITE ELEMENT METHODS FOR SOLVING THE DIFFUSION-CONVECTION-REACTION EQUATION

Ramon Codina

Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports
Universitat Politècnica de Catalunya
Gran Capità s/n, Edifici C1, 08034 Barcelona, Spain

CONTENTS

Submitted for publication to the journal
Computer Methods in Applied Mechanics and Engineering

COMPARISON OF SOME FINITE ELEMENT METHODS FOR SOLVING THE DIFFUSION-CONVECTION-REACTION EQUATION

Ramon Codina

Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports
Universitat Politècnica de Catalunya
Gran Capità s/n, Edifici C1, 08034 Barcelona, Spain

Summary

In this paper we describe several finite element methods for solving the diffusion-convection-reaction equation. None of them is new, although the presentation is non-standard in an effort to emphasize the similarities and differences between them. In particular, it is shown that the classical SUPG method is very similar to an explicit version of the Characteristic-Galerkin method, whereas the Taylor-Galerkin method has a stabilization effect similar to a subgrid scale model, which is in turn related to the introduction of bubble functions.

1 Introduction

The objective of this paper is to compare several finite element methods for solving the linear diffusion-convection-reaction equation from the point of view of the formulation of the methods, describing the motivations that lead to them. Some of these methods have the transient problem as starting point, whereas the others are developed by considering first the stationary equations. Although none of them takes into account whether there is a reaction term in the equations or not, this will lead to an important difference between the methods, as we shall see. This 'reaction' term will be simply a term proportional to the unknown, thus having in fact the physical meaning of absorption for scalar equations.

The methods that will be described and the acronyms that will be used to refer to them are the following:

- SUPG: Streamline-upwind/Petrov-Galerkin method [1].
- ST-GLS: Space-time Galerkin/least-squares method [2].
- SGS: Subgrid scale methods [3–5].
- CG: Characteristic Galerkin method [6–9].
- TG: Taylor-Galerkin method [10].

Essentially, all these methods consist in the addition of a stabilizing term to the original

Although most of the methods to be described in this paper start from the scalar equation (1), we will be also interested in its vector counterpart, that we write as

$$\frac{\partial \boldsymbol{U}}{\partial t} + \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i \boldsymbol{U}\right) - \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{U}}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{U} = \boldsymbol{F} \qquad \text{in } \Omega,\ t \in (0,T), \tag{4}$$

where now $\boldsymbol{U}$ and $\boldsymbol{F}$ are vectors of n_{unk} unknowns and $\boldsymbol{A}_i$, $\boldsymbol{K}_{ij}$ and $\boldsymbol{S}$ are $n_{\text{unk}} \times n_{\text{unk}}$ matrices ($i,j = 1,...,n_{\text{sd}}$). The usual summation convention is implied in Eq. (4).

To simplify the discussion, we assume that the diffusion matrices $\boldsymbol{K}_{ij}$ verify $\boldsymbol{K}_{ij} = \boldsymbol{K}_{ji}^{\text{t}}$ and the bilinear form $\boldsymbol{X}_i \boldsymbol{K}_{ij} \boldsymbol{Y}_j$, with $\boldsymbol{X}_i$ and $\boldsymbol{Y}_j$ vectors of n_{unk} components, is positive definite. Matrices $\boldsymbol{A}_i$ and $\boldsymbol{S}$ are not necessarily symmetric, but it is assumed that there is a matrix $\boldsymbol{T}$ associated to a linear change of variables such that $\hat{\boldsymbol{A}}_i = \boldsymbol{T}\boldsymbol{A}_i\boldsymbol{T}^{-1}$ are symmetric and if $\hat{\boldsymbol{S}} = \boldsymbol{T}\boldsymbol{S}\boldsymbol{T}^{-1}$ then $\frac{\partial \hat{\boldsymbol{A}}_i}{\partial x_i} + \hat{\boldsymbol{S}} + \hat{\boldsymbol{S}}^{\text{t}}$ is positive semi-definite. As in the scalar case, we consider only homogeneous Dirichlet conditions. Under all these conditions, the problem is well posed.

It will be useful in what follows to introduce the following notation:

$$\mathcal{L}_{\text{conv,c}}(\boldsymbol{U}) := \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i \boldsymbol{U}\right), \tag{5}$$

$$\mathcal{L}_{\text{ds}}(\boldsymbol{U}) := -\frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{U}}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{U}, \tag{6}$$

$$\mathcal{L}(\boldsymbol{U}) := \mathcal{L}_{\text{conv,c}}(\boldsymbol{U}) + \mathcal{L}_{\text{ds}}(\boldsymbol{U}). \tag{7}$$

Equation (4) can now be written as

$$\frac{\partial \boldsymbol{U}}{\partial t} + \mathcal{L}(\boldsymbol{U}) = \boldsymbol{F}. \tag{8}$$

Instead of writing Eq. (4) in the *conservative* or *divergence* form, we will be also interested in what we call the *non conservative* version, in which the convective operator $\mathcal{L}_{\text{conv,c}}(\boldsymbol{U})$ is replaced by

$$\mathcal{L}_{\text{conv,nc}}(\boldsymbol{U}) := \boldsymbol{A}_i \frac{\partial \boldsymbol{U}}{\partial x_i}. \tag{9}$$

For the linear equation that we consider, both operators in Eqs. (5) and (9) give the same equation if using the latter matrix $\boldsymbol{S}$ is redefined as

$$\boldsymbol{S} \leftarrow \boldsymbol{S} + \frac{\partial \boldsymbol{A}_i}{\partial x_i}. \tag{10}$$

However, in the applications the expressions (5) or (9) for the convective term come from different forms of the differential equations to be solved (in other words, matrices $\boldsymbol{A}_i$ in (5) and (9) are different). These may be equivalent at the differential level (i.e., for smooth solutions), but in the case of nonlinear problems they may have different weak solutions, a point of special relevance when the numerical solution of these equations is considered. The importance of using the conservative form of the equations is well known in the literature (see e.g. Ref. [21]). We shall comment on the conservation properties of the different schemes to be considered. Unless otherwise stated, we shall always consider that the equation is written using the conservative operator (5). Nevertheless, the operator in Eq. (9) will be useful for presenting the different methods.

$$r(\boldsymbol{U}_h, \boldsymbol{V}_h) = \sum_{e=1}^{n_{\text{el}}} \int_{\Omega^e} \mathcal{P}^e(\boldsymbol{V}_h)^{\,\text{t}} \boldsymbol{\tau}^e \mathcal{R}^e(\boldsymbol{U}_h)\, \mathrm{d}\Omega \tag{17}$$

where $\boldsymbol{\tau}$ is a $n_{\text{unk}} \times n_{\text{unk}}$ matrix of algorithmic parameters with dimensions of time, $\mathcal{P}(\boldsymbol{V}_h)$ is a certain operator applied to the test function and $\mathcal{R}(\boldsymbol{U}_h)$ is a residual of the differential equation to be solved. All these terms will be specified later on for each particular method. The superscript e in Eq. (17) has been used to indicate that the terms are evaluated elementwise (when space-time finite elements are used, these elements have to be considered also in the space-time domain). This superscript will be omitted in the description of the different methods. We shall write also

$$\int_{\Omega'} := \sum_{e=1}^{n_{\text{el}}} \int_{\Omega^e} . \tag{18}$$

2.4 Conservation properties

Let us consider the case in which $\boldsymbol{F} = \boldsymbol{0}$ (and thus $l(\boldsymbol{V}) = 0$) and $\boldsymbol{S} = \boldsymbol{0}$. Suppose also that no boundary conditions are prescribed on $\partial\Omega$. If we define the advective and diffusive fluxes of the unknown $\boldsymbol{U}$ as

$$\boldsymbol{F}_i^{\text{adv}} := \boldsymbol{A}_i \boldsymbol{U}, \qquad \boldsymbol{F}_i^{\text{diff}} := -\boldsymbol{K}_{ij} \frac{\partial \boldsymbol{U}}{\partial x_j}, \tag{19}$$

the bilinear form that should appear in Eq. (12) is

$$a(\boldsymbol{U}, \boldsymbol{V}) = \int_\Omega \boldsymbol{V}^{\,\text{t}} \frac{\partial}{\partial x_i} \boldsymbol{F}_i^{\text{adv}}\, \mathrm{d}\Omega - \int_\Omega \frac{\partial \boldsymbol{V}^{\,\text{t}}}{\partial x_i} \boldsymbol{F}_i^{\text{diff}}\, \mathrm{d}\Omega + \int_{\partial\Omega} \boldsymbol{V}^{\,\text{t}} n_i \boldsymbol{F}_i^{\text{diff}}\, \mathrm{d}\Gamma, \tag{20}$$

where n_i is the i-th component of the external unit normal $\boldsymbol{n}$. If we take $\boldsymbol{V}$ constant in Ω and use the divergence theorem for the first integral in Eq. (20), this reduces to

$$a(\boldsymbol{U}, \boldsymbol{V}) = \boldsymbol{V}^{\,\text{t}} \int_{\partial\Omega} n_i \left(\boldsymbol{F}_i^{\text{adv}} + \boldsymbol{F}_i^{\text{diff}} \right) \mathrm{d}\Gamma, \tag{21}$$

and the integral statement of Eq. (4) reads now

$$\boldsymbol{V}^{\,\text{t}} \left[\int_\Omega \frac{\partial \boldsymbol{U}}{\partial t}\, \mathrm{d}\Omega + \int_{\partial\Omega} n_i \left(\boldsymbol{F}_i^{\text{adv}} + \boldsymbol{F}_i^{\text{diff}} \right) \mathrm{d}\Gamma \right] = 0, \tag{22}$$

for all constant vectors $\boldsymbol{V}$. This implies that the bracketed term must be zero, which is a global conservation statement for the unknown $\boldsymbol{U}$, its fluxes being $\boldsymbol{F}_i^{\text{adv}}$ and $\boldsymbol{F}_i^{\text{diff}}$.

Obviously, the discrete problem will inherit this property provided that the stabilizing term defined in Eq. (17) verifies

$$r(\boldsymbol{U}_h, \boldsymbol{V}_h) = 0, \tag{23}$$

The stabilizing term introduced by the SUPG method has the form given by Eq. (17). where now $\mathcal{P}(v_h)$ is the non conservation form of the convective operator applied to the test function and $\mathcal{R}(u_h)$ is the residual of Eq. (26), that is,

$$\mathcal{P}_{\text{SUPG}}(v_h) = \mathcal{L}_{\text{conv,nc}}(v_h) = \boldsymbol{a} \cdot \nabla v_h, \tag{28a}$$

$$\begin{aligned}\mathcal{R}_{\text{SUPG}}(u_h) &= \frac{\Delta u_h^n}{\Delta t} + \mathcal{L}(u_h^{n+\theta}) - f \\ &= \frac{\Delta u_h^n}{\Delta t} + \nabla \cdot (\boldsymbol{a} u_h^{n+\theta} - k \nabla u_h^{n+\theta}) + s u_h^{n+\theta} - f.\end{aligned} \tag{28b}$$

The SUPG method is consistent, in the sense that the stabilizing term given by Eq. (17) with $\mathcal{R}(u_h)$ defined in Eq. (28b) is zero if u_h is the solution of the continuous (in space) equation (26).

It remains to define the algorithmic parameter τ, which is often called 'intrinsic time'. The way in which it was originally computed goes back to the original idea of the SUPG method, that is, to add numerical diffusion. For that, let us consider the simple one-dimensional model equation:

$$a\frac{\mathrm{d}u}{\mathrm{d}x} - k\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = 0, \qquad 0 < x < 1, \tag{29}$$

with $u(0)$ and $u(1)$ given. If the partition of $[0,1]$ is uniform, h being the element size, and linear elements are employed, it can be shown that the numerical solution is nodally exact if

$$\tau = \frac{\alpha h}{2a}, \tag{30}$$

where

$$\alpha(\mathrm{Pe}) = \coth(\mathrm{Pe}) - \frac{1}{\mathrm{Pe}}, \quad \mathrm{Pe} := \frac{ah}{2k}. \tag{31}$$

In Eq. (31), Pe is the so called (cell) Péclet number and α is the upwind function (a different expression for α is obtained when quadratic elements are used [16]).

In the general case, the strategy usually adopted is to compute τ in Eq. (17) using the straightforward extension from the 1D case, perhaps with slight 'ad hoc' modifications to improve the accuracy in time [ ]. More recently, other ways of computing τ have been proposed on the basis of the convergence analysis of the method, although this has been done mainly for the method to be described in the next section.

The extension of the SUPG method to the vector equation (4) is obvious, except for the definition of τ, that now is a matrix of algorithmic parameters. The expressions of $\mathcal{P}$ and $\mathcal{R}$ are

$$\begin{aligned}\mathcal{P}_{\text{SUPG}}(\boldsymbol{V}_h) &= \mathcal{L}_{\text{conv,nc}}(\boldsymbol{V}_h) = \boldsymbol{A}_i \frac{\partial \boldsymbol{V}_h}{\partial x_i}, \\ \mathcal{R}_{\text{SUPG}}(\boldsymbol{U}_h) &= \frac{\Delta \boldsymbol{U}_h^n}{\Delta t} + \mathcal{L}(\boldsymbol{U}_h^{n+\theta}) - \boldsymbol{F} \\ &= \frac{\Delta \boldsymbol{U}_h^n}{\Delta t} + \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i \boldsymbol{U}_h^{n+\theta}\right) - \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{U}_h^{n+\theta}}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{U}_h^{n+\theta} - \boldsymbol{F}\end{aligned} \tag{32}$$

4 The space-time Galerkin/least-squares method

In the previous section we have described the SUPG method as a finite element formulation to discretize in space a partial differential equation, assuming that the temporal discretization has been already carried out. In particular, we have considered that this discretization has been done using the generalized trapezoidal rule.

Although the approach described above is very common in practical computations, Johnson *et al.* proposed to use the SUPG method together with a finite element discretization also in space [18], based on the use of the discontinuous Galerkin method in time introduced by Lesaint & Raviart for the space discretization of transport equations [19]. The idea behind this was to be able to treat the temporal derivative like the first spatial derivatives.

On the other hand, Hughes *et al.* found [20] that if in the classical Stokes problem for an incompressible fluid the pressure gradient is viewed as a 'convective term' and a SUPG-like strategy is employed for it, it is possible to avoid the need for using different finite element interpolations for the velocity and the pressure satisfying the so called Babuška-Brezzi stability condition (see, e.g., Ref. [21]), which is needed if the standard Galerkin approach is employed. The method proposed first was based on perturbing the original velocity test function of the Galerkin method with a term proportional to the gradient of the pressure test function. The next step was to consider all the Stokes operator applied to the test functions as perturbing term [22, 23]. Going back to the convection-diffusion equation, this idea led to the so called Galerkin/least-squares (GLS) method [2], which is naturally used together with the space-time approach described earlier [24]. In what follows, we consider this space-time method as the natural extension of the GLS method for the steady-state problem, and we refer to them as the space-time Galerkin/least-squares (ST-GLS) formulation.

Before writing down the equations for the ST-GLS method, let us apply the discontinuous Galerkin (DG) method to Eq. (1). To simplify the notation, we consider first the scalar equation, although the following ideas can be directly applied to systems of equations.

Let $t^n = n\Delta t$ and $I^n = [t^n, t^{n+1}]$. The idea of the DG method is to discretize an integral form of the problem to be solved in the space time slab $Q^n = \Omega \times I^n$, enforcing weakly the continuity of the unknown function at time t^n. Both this unknown function and the test functions are allowed to be discontinuous between different space time slabs. Moreover, these can be discretized using completely independent finite element partitions. The simplest way to construct the element domains is to discretize Ω and to take the elements of Q^n of the form $\Omega^e \times I^n$, where Ω^e is an element of the partition of Ω. However, there is no need at all to consider elements prismatic in time.

Let us denote by v^n_+ the upper limit as $t \to t^n$ of a function v of time and by v^n_- the lower limit. The weak form of Eq. (1) in the space time slab Q^n enforcing weakly the continuity condition $u^n_+ = u^n_-$ for the finite element approximation u_h leads to

$$\begin{aligned}\int_{Q^n} \left[v_h \frac{\partial u_h}{\partial t} + v_h \nabla \cdot (\boldsymbol{a} u_h) + k \nabla v_h \cdot \nabla u_h + s v_h u_h \right] \mathrm{d}\Omega \mathrm{d}t \\ + \int_\Omega v^n_{h,+} \left(u^n_{h,+} - u^n_{h,-} \right) \mathrm{d}\Omega = \int_{Q^n} v_h f \, \mathrm{d}\Omega \mathrm{d}t.\end{aligned} \tag{37}$$

This equation must hold for all the test functions v_h defined in the time slab Q^n. If we use the definition of the bilinear form a given in Eq. (14) (now for the scalar case), Eq. (37)

$$
\begin{aligned}
\mathcal{P}_{\text{ST-GLS}}(\boldsymbol{V}_h) &= \frac{\partial \boldsymbol{V}_h}{\partial t} + \mathcal{L}(\boldsymbol{V}_h) \\
&= \frac{\partial \boldsymbol{V}_h}{\partial t} + \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i \boldsymbol{V}_h\right) - \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{V}_h}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{V}_h \\
\mathcal{R}_{\text{ST-GLS}}(\boldsymbol{U}_h) &= \frac{\partial \boldsymbol{U}_h}{\partial t} + \mathcal{L}(\boldsymbol{U}_h) - \boldsymbol{F} \\
&= \frac{\partial \boldsymbol{U}_h}{\partial t} + \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i \boldsymbol{U}_h\right) - \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{U}_h}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{U}_h - \boldsymbol{F}
\end{aligned}
\tag{42}
$$

Remarks

1. Observe that the ST-GLS method does *not* satisfy condition (24) unless $\partial \boldsymbol{A}_i/\partial x_i = \boldsymbol{0}$ and therefore the global conservation property (22) will not hold in general.
2. For comparison purposes, it is interesting to obtain the equations for the ST-GLS method in the particular case of constant-in-time interpolation. In this case, let us write

$$
u_h^{n+1} \equiv u_{h,+}^n = u_{h,-}^{n+1}, \tag{43}
$$

and similarly for the test functions, for which we omit the superscript since they will be the same for all time intervals. The discontinuous Galerkin method given by Eq. (38) now reduces to

$$
\Delta t\, a(u_h, v_h) + \left(v_{h,+}^n, u_{h,+}^n - u_{h,-}^n\right) = \int_{I^n} (v_h, f)\mathrm{d}t. \tag{44}
$$

The left-hand side of this equation may be written as

$$
\int_{Q^n} v_h f \,\mathrm{d}\Omega\mathrm{d}t = \Delta t \int_\Omega v_h \left(\frac{1}{\Delta t}\int_{I^n} f \,\mathrm{d}t\right)\,\mathrm{d}\Omega. \tag{45}
$$

If we consider f time dependent and *continuous in time*, in Eq. (27) we have that

$$
l(v_h) = \int_\Omega v_h f^{n+\theta}\,\mathrm{d}\Omega, \tag{46}
$$

from where it is seen that Eq. (44) is the same as Eq. (27) for $\theta = 1$ but taking an average of f in the time interval I^n instead of the value at time t^{n+1}.
3. Again in the case of constant-in-time interpolation, the terms $\mathcal{P}$ and $\mathcal{R}$ given in Eq. (39) now reduce to

$$
\mathcal{P}_{\text{ST-GLS}}(v_h) = \mathcal{L}(v_h), \tag{47a}
$$

$$
\mathcal{R}_{\text{ST-GLS}}(u_h) = \mathcal{L}(u_h) - f. \tag{47b}
$$

Since the time derivative of u_h does not appear in $\mathcal{R}$, the ST-GLS method will not modify the mass matrix resulting from the DG method given by Eq. (44). This is an important difference between this method and the SUPG method as presented in

convergence analysis of the method with $\mathcal{P}$ given in Eq. (51) and $\nabla \cdot \boldsymbol{a} = 0$ can be found in Ref. [28].

The SGS methods presented in Refs. [3, 4] are a generalization of the above stabilization procedure, which, as we shall see, can be recovered as a particular case. Also particular cases are the stabilization methods based on the introduction of bubble functions to the finite element space. They first attracted interest since it was recognized that the GLS method for the Stokes problem using linear elements is equivalent (up to the choice of the algorithmic parameters) to the use of the Galerkin method with linear elements enriched with bubble functions [29, 30], which are known to be stable [31]. This connection was later on exploited by several authors (see, e.g., Ref. [32]), who proposed different stabilization procedures based on the use of different bubble functions. However, these techniques are related with the use of the term $\mathcal{P}$ given in Eq. (51), and not with the GLS method, as it was pointed out in Ref. [5].

Let us describe now the idea of the SGS methods presented by Hughes in Refs. [3, 4]. Suppose that the unknown u is split as $u = \bar{u} + u'$, where $\bar{u}$ is the part of u which can be represented by the finite element mesh, whereas u' accounts for the unresolvable scales of u, that is, for the variations of u that can not be reproduced because of the mesh size. For example, $\bar{u}$ may be defined as the component of u in the finite element space and u' its component in the orthogonal complement (with respect to a certain inner product) in $\mathcal{W}$.

The strong assumption of what follows is that we assume that u' vanishes on the boundaries of the elements, that is, $u' = 0$ on $\partial\Omega^e$ for $e = 1, 2, ..., n_{\text{el}}$. In this case, u' is the solution of the problem

$$\begin{aligned} \mathcal{L}(u') &= f - \mathcal{L}(\bar{u}) \qquad \text{in} \quad \Omega^e, \\ u' &= 0 \qquad\qquad\quad \text{on} \quad \partial\Omega^e, \end{aligned} \tag{52}$$

which can be solved for u' in terms of the resolvable scale $\bar{u}$ and the Green's function g for the operator $\mathcal{L}$. This leads to

$$u'(\boldsymbol{y}) = -\int_{\Omega^e} g(\boldsymbol{x}, \boldsymbol{y}) \left(\mathcal{L}(\bar{u}) - f\right)(\boldsymbol{x}) \, \mathrm{d}\Omega_x =: M\left(\mathcal{L}(\bar{u}) - f\right)(\boldsymbol{y}), \tag{53}$$

where M is an integral operator and the integral is defined in Eq. (18).

Let us split also the test function v as $v = \bar{v} + v'$. The problem for the resolvable scale $\bar{u}$ is

$$a(\bar{u}, \bar{v}) + a(u', \bar{v}) = l(\bar{v}). \tag{54}$$

Since u' is assumed to vanish on the boundaries of the elements, we have that

$$a(u', \bar{v}) = (\mathcal{L}(u'), \bar{v}) = (\mathcal{L}^*(\bar{v}), u'), \tag{55}$$

where the integral in the L^2 product is again that defined in Eq. (18). Inserting the expression for u' in Eq. (53) into Eq. (55) and using this in Eq. (54) we find that

$$a(\bar{u}, \bar{v}) + \left(\mathcal{L}^*(\bar{v}), M\left(\mathcal{L}(\bar{u}) - f\right)\right) = l(\bar{v}). \tag{56}$$

Observe that up to this point we haven't considered any numerical approximation, that is, Eq. (56) is exact up to the assumption that $u' = 0$ on $\partial\Omega^e$.

$$
\begin{aligned}
\mathcal{P}_{\text{SGS}}(\boldsymbol{V}_h) &= \frac{\partial \boldsymbol{V}_h}{\partial t} - \mathcal{L}^*(\boldsymbol{V}_h) \\
&= \frac{\partial \boldsymbol{V}_h}{\partial t} + \boldsymbol{A}_i^{\mathrm{t}} \frac{\partial \boldsymbol{V}_h}{\partial x_i} + \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{V}_h}{\partial x_j}\right) - \boldsymbol{S}^{\mathrm{t}}\boldsymbol{V}_h \\
\mathcal{R}_{\text{SGS}}(\boldsymbol{U}_h) &= \frac{\partial \boldsymbol{U}_h}{\partial t} + \mathcal{L}(\boldsymbol{U}_h) - \boldsymbol{F} \\
&= \frac{\partial \boldsymbol{U}_h}{\partial t} + \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i \boldsymbol{U}_h\right) - \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{U}_h}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{U}_h - \boldsymbol{F}
\end{aligned}
\tag{63}
$$

Matrix $\boldsymbol{\tau}$ can be defined again as in Eq. (62), now with g being a matrix.

This is the version of the SGS method that we consider in our comparisons, although it is only a particular SGS model. Another possibility is described below.

Remark

If in the original equation (4) the conservative form of the convective term given by Eq. (5) is replaced by the non conservative one in Eq. (9), the perturbation $\mathcal{P}$ for the GLS and the SGS methods become

$$
\mathcal{P}_{\text{GLS}}(\boldsymbol{V}_h) = \frac{\partial \boldsymbol{V}_h}{\partial t} + \boldsymbol{A}_i \frac{\partial \boldsymbol{V}_h}{\partial x_i} - \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{V}_h}{\partial x_j}\right) + \boldsymbol{S}\boldsymbol{V}_h, \tag{64a}
$$

$$
\mathcal{P}_{\text{SGS}}(\boldsymbol{V}_h) = \frac{\partial \boldsymbol{V}_h}{\partial t} + \frac{\partial}{\partial x_i}\left(\boldsymbol{A}_i^{\mathrm{t}} \boldsymbol{V}_h\right) + \frac{\partial}{\partial x_i}\left(\boldsymbol{K}_{ij}\frac{\partial \boldsymbol{V}_h}{\partial x_j}\right) - \boldsymbol{S}^{\mathrm{t}}\boldsymbol{V}_h. \tag{64b}
$$

From this we see that the GLS method satisfies the conservation property (24) when the non conservation form of the equation is used, whereas the SGS is 'consistent', in the sense that condition (24) is fulfilled when the equations are written in conservation form and it is not when the non conservation form is used.

5.3 Approximation to M through bubble functions

In this case, instead of constructing $\check{M}_h$ through an approximation to the Green's function g, we consider directly the finite element approximation to the unresolvable scales u'. The continuous problem of which this function is solution is

$$
a(\bar{u}, v') + a(u', v') = l(v'), \tag{65}
$$

which is nothing but the variational formulation of problem (52).

The function u' can now be approximated by using bubble functions as

$$
u'(\boldsymbol{x}) \approx u'_h(\boldsymbol{x}) = \sum_{j=1}^{n_{\text{bub}}} \psi_j(\boldsymbol{x}) u'_{h,j}, \tag{66}
$$

where n_{bub} is the number of bubble functions ψ and $u'_{h,j}$ are the nodal values of u'_h. Observe that this function satisfies the original assumption of being zero on the element boundaries.

is understood that $\boldsymbol{X}$ depends also on t_{ref} and $\boldsymbol{x}_{\text{ref}}$ through the initial condition (70b). We have that

$$\frac{\mathrm{d}}{\mathrm{d}t}u(\boldsymbol{X}(t),t)|_{t=t_{\text{ref}}} = \left(\frac{\partial u}{\partial t} + \boldsymbol{a}\cdot\nabla u\right)\Big|_{\boldsymbol{x}=\boldsymbol{x}_{\text{ref}},t=t_{\text{ref}}}. \tag{71}$$

If we write the convective term in Eq. (1) as $\boldsymbol{a}\cdot\nabla u + (\nabla\cdot\boldsymbol{a})u$ and use the redefinition (10) (now for the scalar case), Eq. (1) may be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}u(\boldsymbol{X}(t),t) + \mathcal{L}_{\text{ds}}(u)(\boldsymbol{X}(t),t) = f(\boldsymbol{X}(t)), \tag{72}$$

where we have stressed the fact that all the terms are evaluated at $\boldsymbol{x} = \boldsymbol{X}(t)$. The idea now is to discretize the derivative $\mathrm{d}/\mathrm{d}t$ using a finite difference scheme, that is, to discretize the total derivative in Eq. (1) along the characteristics.

Suppose now that we have the solution at time t^n and we want to compute it at time t^{n+1} using the generalized trapezoidal rule, as in section 3 for the SUPG method. Let t_{ref} be a reference time in $[t^n, t^{n+1}]$. The time discretization of Eq. (72) that we consider is:

$$\begin{aligned}\frac{1}{\Delta t}\left[u(\boldsymbol{X}(t^{n+1}),t^{n+1}) - u(\boldsymbol{X}(t^n),t^n)\right] + \theta\mathcal{L}_{\text{ds}}(u)(\boldsymbol{X}(t^{n+1}),t^{n+1}) \\ + (1-\theta)\mathcal{L}_{\text{ds}}(u)(\boldsymbol{X}(t^n),t^n) = \theta f(\boldsymbol{X}(t^{n+1})) + (1-\theta)f(\boldsymbol{X}(t^n)),\end{aligned} \tag{73}$$

where $\theta \in [0,1]$. Once arrived at this equation there are two possibilities, yielding two different versions of the CG method:

6.2 Interpolation of the unknown along the characteristics

This was the method proposed in Refs. [6, 7]. Suppose that $\theta = 1$ and that $t_{\text{ref}} = t^{n+1}$, and write simply $\boldsymbol{x}$ for $\boldsymbol{x}_{\text{ref}}$. Equation (73) in this case reduces to

$$\frac{1}{\Delta t}\left[u(\boldsymbol{x},t^{n+1}) - u(\boldsymbol{X}(t^n),t^n)\right] + \mathcal{L}_{\text{ds}}(u)(\boldsymbol{x},t^{n+1}) = f(\boldsymbol{x}). \tag{74}$$

We may think of $\boldsymbol{x}_{\text{ref}}$ as the configuration at time t^{n+1}. Therefore, the problem is how to evaluate the term $u(\boldsymbol{X}(t^n),t^n)$. For this, it is necessary first to integrate the equation of the characteristics in order to express $\boldsymbol{X}(t^n)$ in terms of the current configuration. This may be done by using either Eqs. (75) or (77) below, depending on the order of accuracy desired. In general, the result will not coincide with any node of the finite element mesh, that is, $\boldsymbol{X}(t^n)$ will lie within an element. This element must be identified and after this the unknown $u(\boldsymbol{X}(t^n),t^n)$ must be interpolated.

6.3 Local expansion of the unknown along the characteristics

We derive now an explicit expression for $u(\boldsymbol{X}(t^{n+1}),t^{n+1})$ and $u(\boldsymbol{X}(t^n),t^n)$ using a Taylor expansion in the neighborhood of $\boldsymbol{x}_{\text{ref}}$. This will allow us to obtain a semi-discrete system of equations where all the terms will be evaluated at the same point of the same spatial domain, thus avoiding the need of finding $\boldsymbol{X}(t^n)$ as described above. This idea can be found in Refs. [8, 9].

Using Eq. (79) in the discretization of the temporal derivative in Eq. (73) (with $\theta = 1/2$) and Eq. (80) to approximate the rest of the terms evaluated at $\boldsymbol{x} = \boldsymbol{X}(t^n)$ and $t = t^n$ we finally obtain

$$\frac{1}{\Delta t}\left[u^{n+1} - u^n\right] + \boldsymbol{a}^{n+1/2} \cdot \nabla u^n + \mathcal{L}_{\mathrm{ds}}(u^{n+1/2}) - f - \frac{\Delta t}{2}\boldsymbol{a}^n \cdot \nabla\left[\mathcal{L}(u^n) - f\right] = 0. \tag{81}$$

Once this semidiscrete problem has been obtained, we may further approximate values at the time level $n + 1/2$ by values at n, thus obtaining a fully explicit scheme. This involves only an approximation of the temporal argument of the functions.

If last term in Eq. (81) is multiplied by a test function v and the result is integrated by parts it is found that

$$-\frac{\Delta t}{2}\int_\Omega v\boldsymbol{a}^n \cdot \nabla\left[\mathcal{L}(u^n) - f^n\right]\, \mathrm{d}\Omega = \frac{\Delta t}{2}\int_\Omega \nabla \cdot (\boldsymbol{a}^n v)\left[\mathcal{L}(u^n) - f^n\right]\, \mathrm{d}\Omega, \tag{82}$$

where we have made use of the fact that $v = 0$ on $\partial\Omega$. For other boundary conditions constant in time it can be also assumed that $\mathcal{L}(u^n) - f^n = 0$ on $\partial\Omega$.

If now the weak form of Eq. (81) is discretized, it is seen that the contribution due to the use of the Characteristic Galerkin method with respect to the standard Galerkin approach has the general form (17), with

$$\mathcal{P}_{\mathrm{CG}}(v_h) = -\mathcal{L}^*_{\mathrm{conv,nc}}(v_h) = \nabla \cdot (\boldsymbol{a}v_h), \tag{83a}$$
$$\mathcal{R}_{\mathrm{CG}}(u_h) = \mathcal{L}(u_h) - f, \tag{83b}$$

where all the terms are evaluated at time step n. Observe that the integral of the RHS of Eq. (82) has to be understood as the sum of the integrals over the element interiors for the discrete problem, that is, in the sense of Eq. (18).

According to the previous derivation, the numerical parameter τ in this case is $\Delta t/2$, the same for all the elements. However, it is shown in Ref. [33] that if instead of taking $t_{\mathrm{ref}} = t^{n+1}$ we take $t_{\mathrm{ref}} = \gamma t^{n+1} + (1-\gamma)t^n$, then $\tau = \gamma\Delta t/2$. The parameter γ is free: it represents the position on the characteristic at which the total time derivative is discretized. This justifies the use of variable τ's.

From the previous derivation of the CG method it is readily seen that there are other schemes with the same accuracy. Our motivation has been to express as many terms as possible evaluated at time step n, although some terms could be equally evaluated at time step $n + 1$, leading to implicit versions of the CG method.

Once the final equations discretized in time have been obtained, it is possible to change the time step at which some terms are evaluated. This will modify the accuracy of the scheme (and perhaps also its stability), but only in the time variable, not along the characteristics.

Remarks

1. It is interesting to note that if the fully explicit version of the CG method is considered and linear finite elements are used, the critical time step above which the scheme becomes unstable turns out to give a value of $\Delta t/2$ very close to the intrinsic time of the SUPG method given in Eqs. (30) and (31) (see Ref. [34]). Also, if $\boldsymbol{a}$ is divergence free it is seen from Eqs. (28a) and (83a) that $\mathcal{P}_{\mathrm{CG}}(v_h) = \mathcal{P}_{\mathrm{SUPG}}(v_h)$.

matrices $\boldsymbol{A}_i$ in Eq. (4) must be replaced by $\hat{\boldsymbol{A}}_i = \boldsymbol{T}\boldsymbol{A}_i\boldsymbol{T}^{-1}$, and $\mathrm{diag}(\hat{\boldsymbol{A}}_i) \neq \boldsymbol{T}\mathrm{diag}(\boldsymbol{A}_i)\boldsymbol{T}^{-1}$ (except, of course, if $\boldsymbol{A}_i = a_i\boldsymbol{I}$, where $\boldsymbol{I}$ is the $n_{\mathrm{unk}} \times n_{\mathrm{unk}}$ identity matrix).

Concerning matrix $\boldsymbol{\tau}$, in principle it is simply $\tau\boldsymbol{I}$, with $\tau = \Delta t/2$. However, as it was mentioned for the scalar case, the use of variable τ's is justified. Moreover, one can also think of using different τ's for the different equations, thus leading to an expression of $\boldsymbol{\tau}$ similar to that given by Eq. (35). This approach is very common in practice and often justified by the use of local time stepping techniques when the transient evolution is not important [34].

7 The Taylor Galerkin method

The Taylor Galerkin method was first introduced by Donea in Ref. [10] as the finite element counterpart of the Lax-Wendroff scheme for finite difference methods. Here we derive a general version of the explicit form of this formulation for Eq. (8). There are also implicit versions of this method, although they have to be motivated using other reasoning.

Let us consider the following Taylor expansion of the unknown $\boldsymbol{U}$ at time step n:

$$\boldsymbol{U}^{n+1} = \boldsymbol{U}^n + \frac{\partial \boldsymbol{U}^n}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 \boldsymbol{U}^{n+\theta}}{\partial t^2}\Delta t^2 + O(\Delta t^3), \tag{87}$$

where $0 \leq \theta \leq 1$. For the moment, let us take $\theta = 0$ (see remark 3 below). If $\boldsymbol{U}$ satisfies Eq. (8) then

$$\boldsymbol{U}^{n+1} = \boldsymbol{U}^n + [\boldsymbol{F}^n - \mathcal{L}(\boldsymbol{U}^n)]\Delta t + \frac{1}{2}\left[\frac{\partial \boldsymbol{F}^n}{\partial t} - \frac{\partial}{\partial t}\left(\mathcal{L}(\boldsymbol{U})\right)^n\right]\Delta t^2 + O(\Delta t^3). \tag{88}$$

As before, we assume that $\boldsymbol{F}$ is time independent. Otherwise, the term $\partial \boldsymbol{F}/\partial t$ should be kept in what follows. If the solution $\boldsymbol{U}$ of Eq. (8) is sufficiently smooth and the coefficient matrices $\boldsymbol{A}_i$, $\boldsymbol{K}_{ij}$ and $\boldsymbol{S}$ are time independent, we have that

$$\frac{\partial}{\partial t}\left(\mathcal{L}(\boldsymbol{U})\right) = \mathcal{L}\left(\frac{\partial \boldsymbol{U}}{\partial t}\right) = \mathcal{L}\left(\boldsymbol{F} - \mathcal{L}(\boldsymbol{U})\right). \tag{89}$$

Using this in Eq. (88) and neglecting the term $O(\Delta t^3)$ we find the following time discretization of Eq. (8):

$$\frac{\boldsymbol{U}^{n+1} - \boldsymbol{U}^n}{\Delta t} = \boldsymbol{F} - \mathcal{L}(\boldsymbol{U}^n) - \frac{\Delta t}{2}\mathcal{L}\left(\boldsymbol{F} - \mathcal{L}(\boldsymbol{U})\right)^n. \tag{90}$$

If this equation is now multiplied by a test function $\boldsymbol{V}$ and integrated over Ω the last term in the RHS of Eq. (90) leads to

$$\frac{\Delta t}{2}\int_\Omega \boldsymbol{V}^{\mathrm{t}}\mathcal{L}\left(\boldsymbol{F} - \mathcal{L}(\boldsymbol{U})\right)^n \, \mathrm{d}\Omega = \frac{\Delta t}{2}\int_\Omega \mathcal{L}^*(\boldsymbol{V})^{\mathrm{t}}\left(\boldsymbol{F} - \mathcal{L}(\boldsymbol{U})\right)^n \, \mathrm{d}\Omega. \tag{91}$$

From this we see that when the weak form of Eq. (90) is discretized the contribution due to the use of the TG method with respect to the standard Galerkin approach has again the general form (17), now with

8 On the discrete maximum principle for scalar equations

8.1 Background

In this section we shall consider the steady version of the scalar equation (1). To simplify further the discussion we take all the coefficients of the equation constant. Thus, the problem we consider is to find u such that

$$-k\nabla^2 u + \boldsymbol{a}\cdot\nabla u + su = f \qquad \text{in } \Omega, \tag{94a}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{94b}$$

For the continuous problem (94) it is well known that the maximum principle holds, that is, the solution attains its maximum at the boundary when f is non-positive. The boundary condition (94b) can be generalized to $u = u_{\mathrm{d}}$, with the given function u_{d} non-negative. The question is whether this property is inherited by the discrete problem or not.

For problem (94) the SUPG and the CG methods in one hand and the SGS and the TG methods in the other coincide, except in the definition of the algorithmic parameter τ. This is why we shall only compare the SUPG, the GLS and the SGS methods here. The study of the discrete maximum principle (DMP) will show an important difference in the behavior of these three methods.

Let n_{tp} be the total number of nodes of the finite element mesh and n_{fp} the number of interior nodes. The finite element discretization of the problem will lead to an algebraic system of the form

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{95}$$

where $\boldsymbol{x}$ stands for the vector containing the nodal unknowns $x_i,\ i = 1, ..., n_{\mathrm{tp}}$. The values $x_i,\ i = n_{\mathrm{fp}}+1, ..., n_{\mathrm{tp}}$ are known from the Dirichlet boundary conditions. Matrix $\boldsymbol{A}$, whose components will be denoted a_{ij}, will have dimensions $n_{\mathrm{fp}} \times n_{\mathrm{tp}}$ and the vector $\boldsymbol{b}$ coming from the source term will have components $b_i,\ i = 1, ..., n_{\mathrm{fp}}$.

As shown in Ref. [35] for linear elements, the satisfaction of the DMP, viz.,

$$\max_{i=1,...,n_{\mathrm{tp}}} \{x_i\} = x_m, \quad \text{with } n_{\mathrm{fp}} + 1 \leq m \leq n_{\mathrm{tp}}, \tag{96}$$

leads to uniform convergence of the finite element solution. Therefore, no spurious oscillations will appear, not even in the vicinity of sharp layers. On the other hand, the DMP follows (see, e.g. Ref. [36]) if $b_i \leq 0,\ i = 1, ..., n_{\mathrm{fp}}$, and matrix $\boldsymbol{A}$ in (95) is of non-negative type, that is,

$$a_{ij} \leq 0 \quad \text{for} \quad i \neq j, \quad i = 1, ..., n_{\mathrm{fp}},\ j = 1, ..., n_{\mathrm{tp}}, \tag{97a}$$

$$\sum_{j=1}^{n_{\mathrm{tp}}} a_{ij} \geq 0, \quad i = 1, ..., n_{\mathrm{fp}}. \tag{97b}$$

Since the assembly operator is nothing but the adequate sum of the element contributions, it suffices to check conditions (97a) and (97b) for the element matrices, hereafter denoted by $\boldsymbol{A}^{(e)}$. Let us split them into their diffusive, convective and reactive contributions as

$$\boldsymbol{A}^{(e)} = \boldsymbol{A}_d^{(e)} + \boldsymbol{A}_c^{(e)} + \boldsymbol{A}_r^{(e)}. \tag{98}$$

It is easy to see that the application of the SUPG, the GLS or the SGS methods to problem (102) is equivalent to the use of the Galerkin method with modified values of the parameters k, a and s. These values are

$$\bar{k} = k + \tau a^2, \tag{105a}$$
$$\bar{a} = a - (\xi + 1)\tau a s, \tag{105b}$$
$$\bar{s} = s - \xi \tau s^2, \tag{105c}$$

where $\xi = 0$ for the SUPG method, $\xi = -1$ for the GLS method and $\xi = 1$ for the SGS method.

Using these effective values for the coefficients of Eq. (102a), condition (104) reads

$$-\frac{k}{h} \pm \frac{a}{2} + \frac{sh}{6} - \tau \left[\frac{a^2}{h} \mp \frac{\xi + 1}{2} as + \xi \frac{s^2 h}{6} \right] \leq 0. \tag{106}$$

This condition is impossible to fulfill for all values of k, a and s, although it provides information about the behavior of the different methods.

First, let us remark that when $s = 0$ condition (106) reduces to

$$\tau \geq \frac{h}{2a} \left(1 - \frac{1}{\mathrm{Pe}} \right), \tag{107}$$

which is the condition that prevents node to node oscillations.

In the limit case $a = 0$ the situation is different. The SUPG method ($\xi = 0$) does not introduce any modification to the Galerkin method, and therefore it is impossible in general to satisfy condition (106) for $a = 0$. For the GLS method ($\xi = -1$) it is easy to see that (106) implies $\tau < 0$, which is incompatible with the case $a > 0$ that leads to condition (107).

Only the SGS method ($\xi = 1$) behaves well in the case $a = 0$. If we define the dimensionless number

$$\mathrm{Ab} := \frac{sh^2}{2k}, \tag{108}$$

which is a measure of the relative importance of the absorption and diffusion terms, condition (106) yields

$$\tau \geq \frac{1}{s} \left(1 - \frac{3}{\mathrm{Ab}} \right). \tag{109}$$

Although the SGS method allows to satisfy condition (106) when $a = 0$, in the general case this is impossible. To see this, observe that the bracketed term in this inequality can be zero for values that lead to a positive left-hand-side. Nevertheless, the limit cases analyzed above provide useful design criteria for τ. First, it is easy to see that for linear elements and in the case $s = 0$ the parameter τ verifies condition (107) if it computed as indicated in Eq. (41) with $C_1 = 1/3$ and $C_2 = 1$, but also if we take

$$\tau = \frac{h}{2a} \frac{\mathrm{Pe}}{\mathrm{Pe} + 1} = \frac{1}{\frac{4k}{h^2} + \frac{2a}{h}}, \qquad \text{for } s = 0. \tag{110}$$

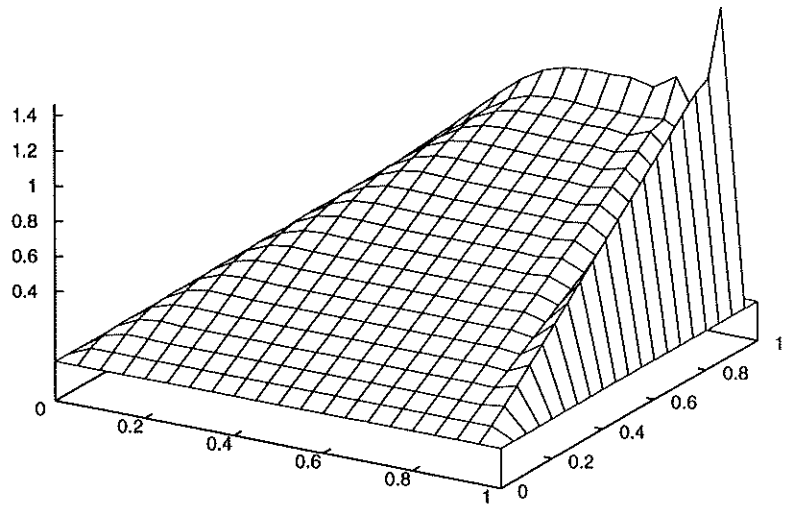

Figure 1. Case $|a| = 1$, $s = 0.0001$, SUPG method.

Results for the first case are shown in Fig. 1. Only those corresponding to the SUPG method have been shown, since for small values of s the GLS and the SGS methods give also the same results. The solution shows some oscillations near the boundary layer created due to the smallness of k.

Results for case b) are shown in Fig. 2. The difference betweeen the three methods is there obvious. Only the SGS method doesn't present any oscillation. The overshoot for the GLS method is stronger that that obtained with the SUPG formulation.

In case c) the effects of convection and reaction are both present, and thus there are oscillations for the three methods due to the presence of convection. Results are shown in Fig. 3. It can be noticed that, even though the convective and reactive terms have a similar influence in the solution for the values of the parameters taken, Ab is much smaller than Pe, that is to say, the oscillations are dominated by those due to convection.

In Fig. 4 we have plotted the results obtained using the SUPG method on a much finer mesh of 52×52 bilinear elements refined near the boundaries. There is only an overshoot at $(1, 1)$ in the cases with convection.

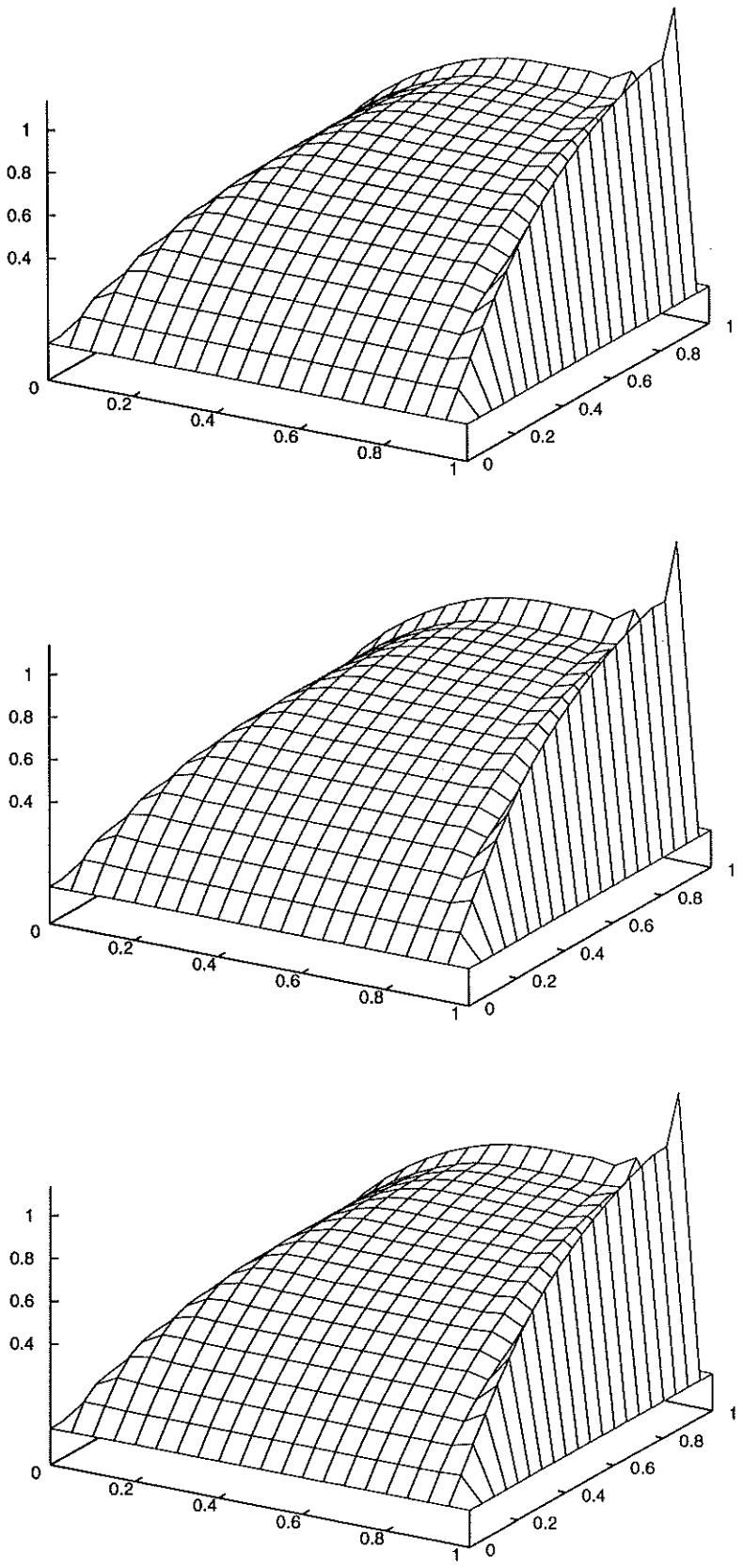

Figure 3. Case $|\boldsymbol{a}| = 0.5$, $s = 1$. From the top to the bottom: SUPG, GLS and SGS methods.

[8] R. Löhner, K. Morgan and O.C. Zienkiewicz, 'The Solution of Non-linear hyperbolic Equation Systems by the Finite Element Method', *Int. J. Num. Meth. Fluids*, **4**, 1043–1063 (1984).

[9] O.C. Zienkiewicz and R. Codina, 'A general algorithm for compressible and incompressible flow. Part I: the split, characteristic based scheme', *Int. J. Num. Meth. Fluids*, **20**, 869–885 (1995).

[10] J. Donea, 'A Taylor-Galerkin Method for Convection Transport Problems', *Int.J. Num. Meth. Engrg*, **20**, 101-119 (1984).

[11] R.J. LeVeque, *Numerical methods for conservation laws*, Birkhäuser, 1990.

[12] J. von Neumann and R.D. Richtmyer, 'A method for the numerical calculation of hydrodynamical shocks', *J. Appl. Phys.*, **21**, 232 (1950).

[13] D.W. Kelly, S. Nakazawa, O.C. Zienkiewicz and J.C. Heinrich, 'A note on upwinding and anisotropic balancing dissipation in finite element approximations to convective diffusion problems', *Int. J. Numer. Meth. Engrg.*, **15**, 1705–1711 (1980).

[14] T.J.R. Hughes and A. Brooks, 'A multi-dimensional upwind scheme with no crosswind diffusion', in: *FEM for convection dominated flows*, T.J.R. Hughes (ed.) ASME, New York (1979).

[15] T.J.R. Hughes and A.N. Brooks, 'A theoretical framework for Petrov-Galerkin methods, with discontinuous weighting functions: applications to the streamline upwind procedure', in: *Finite Element in fluids*, R.H. Gallagher, D.M. Norrie, J.T. Oden and O.C. Zienkiewicz (eds.), vol. IV, (Wiley, London, 1982) 46–65.

[16] R. Codina, E. Oñate and M. Cervera, 'The intrinsic time for the SUPG formulation using quadratic elements', *Comput. Meths. Appl. Mech. Engrg.*, **94**, 239–262 (1992).

[17] T.J.R. Hughes and M. Mallet, 'A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective-diffusive systems', *Comput. Meths. Appl. Mech. Engrg.*, **58**, 305–328 (1986).

[18] C. Johnson, U. Nävert and J. Pitkäranta, 'Finite element methods for linear hyperbolic equations', *Comput. Meths. Appl. Mech. Engrg.*, **45**, 285–312 (1984).

[19] P. Lesaint and P.A. Raviart, 'On a finite element method for solving the neutron transport equation', in: C. de Boor (ed.), *Mathematical aspects of the finite element method*, Academic Press, 1974.

[20] T.J.R. Hughes, L.P. Franca and M. Balestra, 'A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuska-Brezzi condition: a stable Petrov-Galerkin formulation for the Stokes problem accommodating equal-order interpolations', *Comput. Meths. Appl. Mech. Engrg.*, **59**, 85–99 (1986).

[21] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, 1991.

[22] T.J.R. Hughes and L.P. Franca, 'A new finite element formulation for computational fluid dynamics: VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity/pressure spaces', *Comput. Meths. Appl. Mech. Engrg.*, **65**, 85–96 (1987).

[23] L.P. Franca and R. Stenberg, 'Error analysis of some Galerkin least-squares methods for the elasticity equations', *SIAM J. Numer. Anal.*, **28**, 1680–1697 (1991).

[24] F. Shakib and T.J.R. Hughes, 'A new finite element formulation for computational fluid dynamics: IX. Fourier Analysis of space-time Galerkin/least-squares algorithms', *Comput. Meth. Appl. Mech. Engrg.*, **87**, 35–58 (1991).

[25] I. Harari and T.J.R. Hughes, 'What are C and h ? : Inequalities for the analysis and design of finite element methods', *Comput. Meth. Appl. Mech. Engrg.*, **97**, 157–192 (1992).

[26] F. Shakib, T.J.R. Hughes and Z. Johan, 'A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier-Stokes equations', *Comp. Meth. Appl. Mech. Engng.*, **89**, 141–219 (1991).

[27] J. Douglas and J. Wang, 'An absolutely stabilized finite element method for the Stokes problem', *Math. Comput.*, **52**, 495–508, (1989).